

## Improving Middle School Math Outcomes Through Tutoring

### Introduction

In the United States in 2019, 34% of eighth grade student's math scores were on grade level or higher (Mervosh and Wu, 2022). Two years later, after COVID interrupted learning for districts across the country, the number of eighth grade students proficient in math shrunk to 26%. Los Angeles Unified School District (LAUSD) was no exception, decreasing from 33% to 28% (EdSource, 1). LAUSD, the second largest school district in the country, educates a high percentage of historically marginalized students: 75% Hispanic, 10% Black, 45% English Langue Learners (ELLs), and 12% have a documented disability (IEP). While some student demographics within LAUSD have scored much higher than the national average before and after COVID, namely Asians and Whites, other demographics have scored much lower. As shown in figure 1, COVID exacerbated the near 50% difference in grade level math proficiency between Asian and Black and Brown students within LAUSD.

As achievement gaps remain stagnant, or in some instances increase, school districts are left with the challenge of improving student math scores. More funding would certainly help schools with high percentages of low-achieving students by recruiting more experienced teachers and administrators and improving the quality of school resources. However, given the U.S. nascent neoliberalism approach to K-12 education, such government spending seems unfeasible. In fact, many schools will have a smaller budget starting in August 2024 as temporary funds granted during COVID will expire. Schools in LAUSD are scrambling to implement cost-effective ways to improve middle school math outcomes, especially for marginalized students.

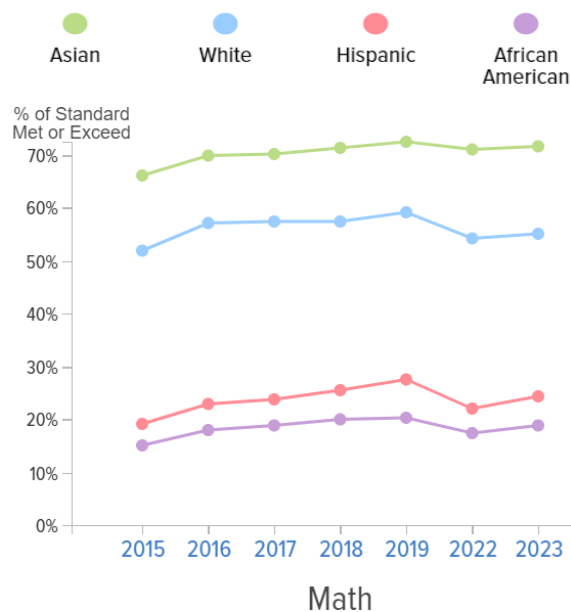
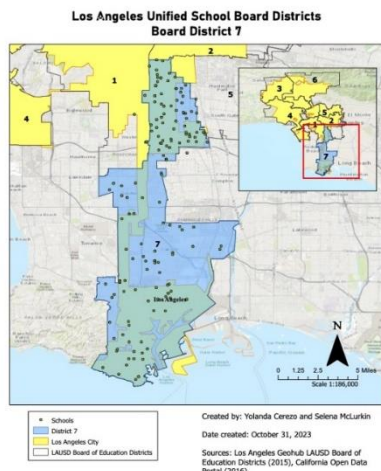


Figure 1

Source: CAASP.Edsource.org



A sub-district within LAUSD, Board District 7 (BD7), encompasses 14 middle schools that serve marginalized, low-performing students. These schools are severely under-resourced and have some of the lowest math scores in the district. As part of my senior capstone project, we are working with BD7 to determine cost-effective ways to improve student math scores. Research has shown that tutoring is highly effective at improving students' academic performance. However, tutoring is a broad term and encompasses an array of various models. This paper will explore tutoring interventions that Board District 7 may learn from and implement. The essay will explore three case studies,

each highlighting a different type of tutoring: (1) peer-peer tutoring, (2) homogenous, small group tutoring, and (3) shadow education.

### *Background on BD7 and Defining Tutoring*

Due to the size of LAUSD, the district maintains a board of seven members, each representing a different district in LAUSD. The board member and their office are responsible for identifying and monitoring student achievement goals, supervising the superintendent, developing, and adopting education policies, influencing budget expenditures, and collaborating with parents to resolve issues in their district. While each board member cannot unilaterally implement policy in sub-district, they can influence the education policies that the board of LAUSD implements. Board District 7, serving over 60,000 students across 175 schools, is currently chaired by Tanya Ortiz Franklin. Franklin was elected to this position by her constituency in 2020 and is running for reelection in 2024.

Tutoring is an incredibly broad term that contains a variety of disparate models, such as students informally teaching other students, computer programs teaching students remedial subjects, heterogeneous small groups, homogenous small groups, pull-out methods, push-in methods, and many more. A common definition across the growing body of research defines tutoring as, “One-on-one or small-group instructional programs” (Nickow et. al, 2020). While the definition is still broad, it does delineate whole-class instruction from small groups which is an important distinction.

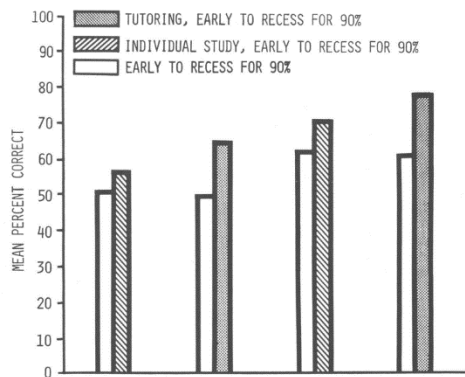
### *Case Study #1: Peer-Peer Tutoring*

This study occurs in an Indian youth center in Arizona, 1973. The student demographics are third through fifth graders and most students were one to two years behind in math, similar to the demographics and student achievement levels of the students in BD7. The researchers examined how unstructured, peer-peer tutoring affected student learning outcomes, particularly lower-performing students. The study defines peer to peer tutoring as, “students intentionally aiding the learning of other students” (Stigmar, 2016).

To examine the effects, the researchers added an intervention of peer-peer tutoring for 15 minutes before student’s math quiz. Students were not given much instruction, other than to, “help classmates with problems on the worksheets” (Harris et. al, 1973). The student instructors were identified by their high performance (greater than 90%) on a prior quiz. There were several control groups: one did not receive any tutoring nor independent study prior the quiz, another group was allowed to study independently for 15 minutes prior to the quiz, and the last group had an incentive of leaving early to recess if students scored higher than 90% on their quiz.

The researchers found that students who received tutoring for just 15 minutes before the test showed the highest increase in their math scores, about 10% (figure 2). The results were even

higher when the peer-to-peer tutoring was coupled with the incentive of early recess for achieving a score higher than 90% (figure 3). Individual study improved average test scores by about 5%. The learning outcomes persisted for multiple months, and student instructors also improved their math scores, although not as much as students receiving tutoring.



Source: (Harris et. al, 1973)

Peer-peer tutoring significantly improved students' scores, was an affordable intervention, and required little teacher training or preparation. While peer-peer tutoring occurred, teachers could work with students who needed the most help, increasing teacher efficiency as they were able to intentionally target help to those most in need.

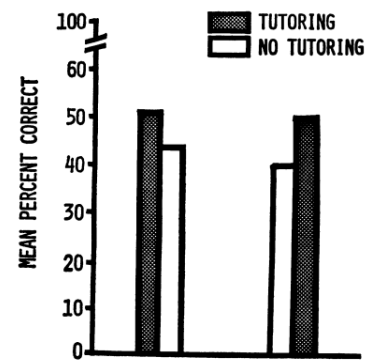


Figure 2

Source: (Harris et. al, 1973)

However, the study did not communicate the education and experience background of the teachers they chose to conduct the study. How many years did those teachers teach? Did they have experience with peer-peer tutoring? How strong were their classroom management skills? All these variables are important in repeating the effectiveness of this peer-peer tutoring program. A teacher with little classroom experience, or who struggles with classroom management would struggle to create a classroom culture where students want to learn and teach each other. Scores would be significantly lower if students were not bought in the classroom, especially given such an informal model. Additionally, the study was conducted fifty years ago which raises questions. Are there generational changes in how students learn which would change the results from this study? How has the influence of technology changed student's attention spans? Are students today incentivized by recess to the same extent that students in 1973 were incentivized? Overall, peer-peer tutoring seems like a viable, cost-effective solution for improving student's math scores in BD7.

#### Case Study #2: Homogenous, Small Groups (Norway)

The second case study is a research project that was funded by the Norwegian government in 2022 in which 159 elementary schools were monitored over four years, totaling 30,000 students. The researchers sought to better understand the effects homogenous, small-group, pullout instruction had on students' math scores. Homogenous grouping places students of similar ability into groups of about three to four students. Each group of students is pulled out for three to four hours per week for a total of eight to twelve weeks (low-dose tutoring), and every student was pulled out, not just low-performing students.

The Norwegian government paid 20 million Euros hire trained, credentialed teachers to lead small groups (Bonesrønning et. al, 2022). Each tutor collaborated with the classroom teacher to mirror the content covered in the classroom and adjusted the material to fit the levels of each group. The tutoring program was implemented in 78 schools, and was not implemented in 81 schools, serving as the control. The researchers found that all students who received tutoring showed growth, regardless of their baseline score, gender, or their family socio-economic status. Specifically, the study found that students improved on Norway's national math test by .06 standard deviations, and that low-performing students had a bigger range of effects than low, middle, and high performing students (figure 4). As shown, low-performing students had the lowest marginal effect and the highest marginal effect compared to medium and high performing groups. On average, middle performing students benefited the most from this program, and low and high performing students benefited to similar levels.

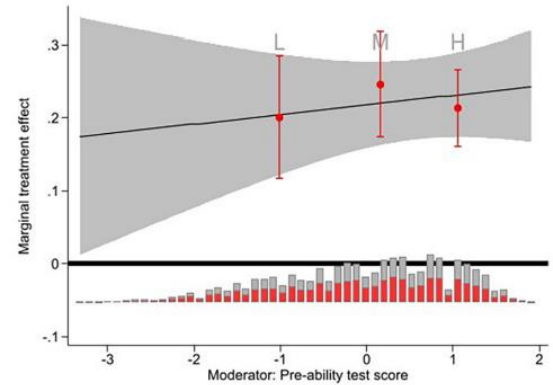


Figure 4

Source: (Bonesrønning et. al), 2022

Relative to the peer-to-peer tutoring model, this program is much harder to implement. For one, Norway's model is highly structured given that teachers are required to collaborate with tutors to plan small-group lessons, monitor student growth, adjust groups, and share successes and challenges. Such collaboration would require teachers to have more planning time, which is a daunting ask for understaffed schools. Furthermore, this program is much more expensive than the peer-peer model. The grant for this program was around 20 million Euros. If schools had extra money to hire staff, it would be more cost-effective to hire more SPED educators, general educators, or campus monitors, instead of hiring tutors. Also, under-resourced schools typically have a high percentage of teachers without a credential or have less than three years of experience. Inexperienced or uncredentialed teachers may not be organized and experienced enough to collaborate with another instructor, provide recommendations on student groups or advice on pedagogy.

Culturally, Norway has a Keynesian education system which has a smaller discrepancy between low and high performing students. Given the small range, Norway can give more attention to each student class, not just low-performing students. However, the United States has a neoliberal K-12 education system which has created a big achievement gap between low and high performing students across the country and in school districts like LAUSD. Given the wide gap, most educators and policy makers focus on low performing students, as the system is already working for the students who are performing well. Thus, a tutoring system that focuses on *all* students is politically unfeasible in a culture like the United States, where extremes between low and high performing are great. Although the study boasts a .06 improvement in the standard deviation of student's national test scores, the number is small. A .06 improvement in standard deviation would hardly close any achievement gaps. Despite being statistically significant, the outcome does not have much pragmatic significance.

Overall, a low dose tutoring model and homogenous are interesting characteristics to consider when considering tutoring models that BD7 might implement. This exact model is too costly,

structured, and culturally specific to Norway to emulate in the United States, however parts of the program could be useful in designing a successful tutoring intervention.

### *Case Study #3: Shadow Education*

Shadow education, private one-on-one tutoring, has existed in eastern Asia for many years as a mechanism for wealthier families to maintain their elite status by ensuring their children receive more education than their peers (Bray, 2019). In the 1990s, shadow education proliferated throughout eastern Europe as the fall of the Soviet Union brought economic hardship upon families who then sought intense tutoring as an investment in their children's prosperity. More recently, shadow education has surged throughout western Europe as high-stakes exams have increased stress on students and families. Shadow education is especially prevalent in middle and upper-income families who feel that traditional education is not enough to ensure their children a secure position in society. Compared to other models of tutoring, shadow education is more of a product of globalization and neoliberalism.

Wei Zhang researchers the shadow education landscape in Chongqing, China, which has a substantial rural and urban population, totaling six million people. Zhang first conducts a literature review on the national cultural context of shadow education before conducting qualitative research (interviews and surveys) on 9<sup>th</sup> grade families and students who live in both rural and urban environments. In terms of supply, inexperienced teachers or teachers who work in rural schools have very low salaries. Through interviews, Zhang discovers that some teachers can double or triple their teacher salary through tutoring, given that private tutoring is part of the informal economy and not taxed (Zhang, 2014). Zhang identifies political and economic policies that increase the demand for shadow education. Recent economic growth has increased the disposable income of middle- and upper-class families. Additionally, China's one-child policy has decreased the number of offspring in families, which has funneled more money to less children (Zhang, 2014). Therefore, families are able and willing to spend a substantial amount of income on shadow education. Zhang explains how two national exams, the senior high school entrance exam (9<sup>th</sup> grade) and the national college entrance exam (12<sup>th</sup> grade), have put incredible stress on students and families as these exams effect student's education and work opportunities. Education policy makers wanting to decrease student's stress, enacted policies that shortened school hours and decreased the amount of homework. However, since the stakes of the exams remained unchanged, students and families felt more unprepared, increasing the demand of shadow education.

After contextualizing the cultural and political forces supporting shadow education, Zhang interviewed and surveyed over 150 students and parents. Since shadow education is part of the informal economy, uncertainty exists around the average pay for teachers, quality of private tutoring, average amount of hours tutored, how parents choose a tutor, and common tutoring topics (Bray, 2019). Zhang's goal in conducting qualitative research is to fill this information gap and contribute to the growing understanding of shadow education. His research found that 45% of students received tutoring. Of those receiving tutoring, 84% of students attended prestigious schools, 74% of students attended urban schools and just 11% of students were enrolled in rural schools (Zhang, 2014). Over ninety percent of students received tutoring from a teacher, as opposed to a tutor, and math was the most common topic that students received tutoring.

The discrepancy in the amount of urban versus rural students receiving shadow education illustrates how private tutoring exacerbates achievement gaps (Zhang, 2014) (Bray, 2019). Urban students and families either feel more stress from these national exams or have more resources (connection to tutors, money) to enroll their student in shadow education. Students and families may feel more pressure to hire a private tutor if other students at their school are enrolled in shadow education. In respect to Board District 7, private tutoring is not a model that a district would implement. However, this research highlights the pervasive effects that testing can have on students, families, and their communities. More research should be conducted on how tests and attendance on school negatively affects low-performing students. Zhang's literature review highlights the caution needed when implementing education policies. Whether politicians are attempting to decrease the effects of stress from national exams, or a district is implementing a tutoring program, students and families needed to be involved in decision making to increase the chances of a policies' effectiveness. More research on shadow education in the United States is needed, especially in low-income, urban areas. Are there inexpensive one-on-one tutoring programs being offered? If so, are they effective? If not, is this a role that education non-profits can occupy?

### *Conclusion*

Of the three tutoring models presented in the case studies, the peer-peer model appears to be the most cost-effective. It's the cheapest, requires the least amount of training, and showed significant improvements to student's math scores. More research is needed to determine characteristics of strong peer-peer programs. Does more than fifteen minutes of tutoring yield substantial returns? Is it best to pair students one to one, or to have a student with high mastery of the material tutor groups of two to three students? How does teacher experience impact the outcomes of peer-peer tutoring? What classroom activities promote a strong classroom culture that encourages positive returns to peer-peer tutoring? These questions should be answered before any policy is recommend to Board District 7.

Tutoring offers a feasible, cost-effective strategy to improve student learning outcomes, however this intervention alone will not close achievement gaps between low and high performing students. The neoliberalist approach to education in the United States bifurcates students into low and high performing groups which places a significant ceiling on careers and earnings. While tutoring improves educational inequities, bottom-up approaches alone will not be enough. Instead, top-down policies should be implemented in concert with bottom-up approaches. More funding, promoting teacher training and credentials, and smaller class sizes would all increase student's mastery of material. However, such top-down approaches are outside the locus of control for educators and given the rising social pressure of increasing achievement gaps, educators are pressured to act. Peer-peer tutoring programs are a viable, cost-effective option to help schools and educators to do what they can.

### *Works Cited*

- Nonearning, Hans, et al. "Small-group instruction to improve student performance in mathematics in early grades: Results from a randomized field experiment." *Journal of Public Economics* 216 (2022): 104765.
- Bray, Mark. "Shadow education in Europe: Growing prevalence, underlying forces, and policy implications." *ECNU Review of education* 4.3 (2021): 442-475.
- Harris, V. William, and James A. Sherman. "Effects of peer tutoring and consequences on the math performance of elementary classroom students 1." *Journal of Applied Behavior Analysis* 6.4 (1973): 587-597.
- "Los Angeles Unified Smarter Balanced Test Results." *EdSource*, [caaspp.edsource.org/sbac/los-angeles-unified-19647330000000](https://caaspp.edsource.org/sbac/los-angeles-unified-19647330000000). Accessed 14 Dec. 2023.
- Martin Stigmar (2016) Peer-to-peer Teaching in Higher Education: A Critical Literature Review, *Mentoring & Tutoring: Partnership in Learning*, 24:2, 124-136, DOI: 10.1080/13611267.2016.1178963
- Mervosh, Sarah, and Ashley Wu. "Math Scores Fell in Nearly Every State, and Reading Dipped on National Exam." *The New York Times*, The New York Times, 24 Oct. 2022, [www.nytimes.com/2022/10/24/us/math-reading-scores-pandemic.html](https://www.nytimes.com/2022/10/24/us/math-reading-scores-pandemic.html).
- Nickow, Andre, et al. "The Impressive Effects of Tutoring on Prek-12 Learning: A Systematic Review and Meta-Analysis of the Experimental Evidence." *NBER*, 6 July 2020, [www.nber.org/papers/w27476](https://www.nber.org/papers/w27476).
- Zhang, Wei. "The demand for shadow education in China: Mainstream teachers and power relations." *Asia Pacific Journal of Education* 34.4 (2014): 436-454.